

Department of Human Services

DHS Data Warehouse Reference Guide



*Making knowledge available when, how,
and where it is needed*

Table of Contents

INTRODUCTION.....	1
OUR MISSION STATEMENT.....	1
GOALS.....	1
ABOUT THE WAREHOUSE.....	2
HOW THE WAREHOUSE WORKS.....	2
RELATIONSHIP OF WAREHOUSE TO TRANSACTION SYSTEMS	2
HOW THE WAREHOUSE IS BUILT AND REFRESHED	2
IMPORTANT WAREHOUSE FACTS.....	3
ACCESS	3
INTERNET ACCESS	3
REQUESTING ACCESS	3
RESPONSIBILITY AND CONFIDENTIALITY	4
QUERYING DATA WITH SECURITY RESTRICTIONS	4
RELEASING DATA OUTSIDE THE DEPARTMENT	4
KEEP A LOG	4
RELEASING DATA WITHIN THE DEPARTMENT.....	4
DESKTOP SECURITY	5
QUERY TOOLS.....	5
THE COGNOS SUITE AND THE COGNOS WEB SUITE	5
THE TOOL YOU SHOULD USE.....	6
OTHER TOOLS	6
TRAINING.....	6
SUPPORT SERVICES.....	7
TECHNICAL SUPPORT	7
DATA WAREHOUSE MAILING LIST.....	7
DATA WAREHOUSE USER GROUPS.....	7
WAREHOUSE UPDATES.....	7
REPORTING DATA QUALITY ISSUES.....	7
WHAT THE WAREHOUSE CONTAINS	8
CLIENT INFORMATION	8
<i>Refresh Cycle</i>	8
<i>Source of the Data</i>	8
<i>Time Span Covered</i>	8
REQUESTING ADDITIONS AND CHANGES TO THE WAREHOUSE	8



DHS Data Warehouse Reference Guide

Introduction

Welcome to the DHS Data Warehouse (Warehouse). The Warehouse allows decision makers at all levels within the department to query, analyze, and report department activities from their web browsers or desktops.

The Warehouse stores information from different information systems, such as:

- USSDS (Unified Social Services Delivery System),
- JIS (Juvenile Information System),
- HRE (Human Resource Enterprise System).

Future sources of information will include the Child Welfare Information System, Medicaid, and others.

Information about clients and services stored in these systems is merged, cleansed, transformed, and loaded into the Warehouse on a weekly, monthly, quarterly, or annual cycle. Demand for information determines frequency of Warehouse updates.

Our Mission Statement

Data Administration manages data as a department resource to enhance client service delivery by making accurate information available to the department community, the public, and other organizations when, how, and where they need it.

Just as other offices manage the department resources of funds, programs, and people, Data Administration manages the resources of department data. Department data is data relevant to planning, managing, operating, auditing, and communicating department activities.

Goals

- Improve the quality of data, including accuracy, timeliness, and definition, for research and decision making.
- Assure that data is kept secure, including confidentiality and protection from loss.
- Improve ease of access, assuring the data is easy to locate, access, and understand.
- Promote flexible, cost effective systems, using existing resources where possible, to answer today's questions.

- Minimize cost to agencies.

About the Warehouse

The warehouse resides on a Sybase database server. Authorized individuals of the DHS community can access the Warehouse through desktop and web-based tools. Access does not require formal programming skills. More information about the Warehouse is available in the following documents:

Tables (www.hsdw.state.ut.us/reports/dwtables.htm)

Notices (www.hsdw.state.ut.us/news/notices.htm)

How the Warehouse Works

The Warehouse is built from operational (or transactional) data, which is then made available for query, analysis and reporting.

Relationship of Warehouse to Transaction Systems

Transaction systems handle the day-to-day functions of a business--recording, processing, and storing operational data. Examples of transaction systems include the USSDS, JIS, HRE, Child Welfare Information System, State Financial Accounting System, and so on. These systems handle business functions such as managing child welfare cases, managing juvenile justice cases, making payments to providers of services, printing a paycheck, or recording the hours worked by employees. These transaction systems, however, were not designed to provide managers with the information they need to conduct trend analyses or to undertake longitudinal studies that require use of historical department data. Data warehouses, unlike transaction systems, are specifically designed to handle these types of analyses.

How the Warehouse is Built and Refreshed

To build the Warehouse, data from different transaction systems was analyzed, cleaned, restructured, supplemented, and summarized to support query, analysis and reporting:

- Analysis involves determining whether each piece of data in the transaction system should be included in the Warehouse.
- Cleaning involves converting old codes into new codes, such as reducing 9 genders to 3 genders, "Male", "Female", and "Unknown".
- Restructuring involves such tasks as breaking certain data elements apart into new data elements. For example, where the name of a person exists in a single field, new fields in the Warehouse would include three separate fields, "Last Name", "First Name", and "Middle Name".
- Supplementing involves adding new data elements. For example, "Date of Birth" and "Service Start Date" both exist as data elements in the Warehouse. We calculate the difference between these two data elements and store a new one, "age at first contact", in the Warehouse. We also supplement the Warehouse by making it easier to use, such as adding a service description, instead of just the service code.
- Summarization involves counting, summing, averaging, etc., certain data elements in the Warehouse. For example, we can count the number of clients receiving service each fiscal year by each service agency.

The Warehouse is refreshed periodically, using the transaction systems as its sources of data. That is, programs extract data from each transaction system and then translate it into Warehouse format. The formatted files are then transferred and loaded into the Warehouse. The data is then available for query, analysis and report.

Note that each subject area has its own refresh cycle governing when its data is periodically updated in the Warehouse. For example, client and service data is refreshed weekly, while payments to providers of services is refreshed monthly. Thus, data in the Warehouse reflects the most recent refresh cycles for the subject areas. Recipients of reports from the Warehouse need to be aware of this "data delay" in the Warehouse. Refresh cycle information is noted in the documentation for subject area under the section What the Warehouse Contains. Refresh cycles are subject to change; Data Administration will notify you of any change in refresh cycles by e-mail.

Important Warehouse Facts

The following are important facts that you should know about the Warehouse:

Data in the Warehouse is for query, analysis, and reporting only. That is, you cannot add, change, or delete data in the Warehouse. These activities must be done only in the transaction system. Once completed there, then the addition, change, or deletion should be included in the next refresh cycle for that subject area in the Warehouse. See Data Quality for more information.

The refresh cycle varies for each subject area in the Warehouse. Check Notices (www.hsdw.state.ut.us/news/notices.htm) to determine the last update for the subject area in which you are interested.

As previously mentioned, there is a "data delay" between the Warehouse and the transaction systems. Keep this time frame in mind when comparing data between the Warehouse and its source systems or when analyzing query results--it may help you to understand and account for any data discrepancies found. Always indicate or communicate the date on which your Warehouse results were obtained and identify the Warehouse, and not the transaction system, as the source of the information. For example, include the date of the query in the report header and the source of the information, the DHS Data Warehouse, in the report footer.

Some data are more dynamic--change more frequently -- than other data, or are subject to change at specific times during the fiscal year. Take this into consideration when writing queries against the Warehouse. Examples of data that change more frequently are clients and services. New clients and services can be added at any time in the source systems.

Access

Internet Access

You can access the Warehouse (www.hsdw.state.ut.us) from your Internet browser as long as you have a password and authorization to use the Warehouse.

Requesting Access

Request authorization and access by contacting any member of the Warehouse Team (www.hsdw.state.ut.us/support.htm).

Responsibility and Confidentiality

The Warehouse contains confidential and sensitive department data. In order to use its data, you must have proper authorization. Your authorization means that you have the authority to use the data and the responsibility to share stewardship of the data with the other users of the Warehouse.

Once authorized, you can access the data that you need to do your job. All authorized users are cautioned, however, that they are entrusted to use the data they retrieve from the Warehouse with care. Confidential data must not be released to others except for those with a "legitimate need to know."

Note also that some data may be restricted by division or cannot be released to persons outside the division. Persons who work with client data should also be aware of any state or federal laws protecting client data.

Querying Data with Security Restrictions

If you execute a query requesting data that you are not authorized to access, you will get results which may be incomplete because they are missing the data you are not allowed to access.

If your authorization is limited to a specific set of data, be sure when querying the data that your record selection conditions include your security restrictions. For example, if you are authorized to access just data for a particular division, one of your record selection conditions should state something like "If Service Agency = 'My Service Agency'," where My Service Agency is your division. This will document why the query gets the results it does, and will also help your query run faster.

Releasing Data Outside the Department

Only summarized or aggregated information may be disclosed outside of the department. This means that there is nothing in the information provided that can identify a specific client. For example, number and type of clients served this year compared to last.

No one may disclose identifying information about clients outside the department, except when it is part of their job responsibilities to do so. For example, providing information about a specific client to a Juvenile Court Judge.

Keep a Log

You might want to keep a log of the reports you create, even if you are giving it to someone within the department. The log might include:

- Who requested the data
- When they made the request
- What data they requested
- Why they requested the data/how they planned to use it
- What query and/or report you used.

Releasing Data Within the Department

Within the department, client data may be disclosed only if it is needed to do the business of the department, and only to those who need to know the information in order to do their jobs.

If you are not sure whether to fulfill a request for client data, call the Office of the Executive Director.

Questions you should ask to help decide whether to fulfill a request for client data:

- Who wants the information?
- Why do they want the information?
- For what purpose will the information be used?
- If they pass the information on to someone else, for what purpose will that person use it?
Note: Usually, the data should be for the requestor's use only. No data should be posted in a public place including the web.
- How will they secure the information once they have it?
- How will they dispose of the information when they are done with it? For example, hard copy reports should be shredded.

Desktop Security

Query Results. Cognos writes reports to the desktop client under the subdirectory specified when the reports are saved. The default directory is C:\Program Files\Cognos\Reports. You must see to it that any sensitive data stored on your personal computer is safeguarded through physical security, access control software, or encryption.

Examples of physical security are locked offices and locked keyboards.

Examples of access control software are a screen saver with password protection (which your computer has been set up to initiate at startup) or specialized desktop security software.

If you encrypt your query results, you will need to decrypt them before accessing them with Cognos.

Warehouse Access. When a computer is left signed on to an account, it is easy for someone to gain unauthorized access. Either sign off from your account before you leave your computer or restrict access by some other means (physical security or access control software).

For more information on security, contact Jerry Smith at 538-4297 or jsmith@hsadmin.hsadm2.state.ut.us.

Query Tools

The Cognos Suite and the Cognos Web Suite

The Cognos Suite and the Cognos Web Suite are the primary supported query, analysis, and reporting tools used to provide access to the Warehouse, which resides on a Sybase database server. Through a point-and-click interface, users can perform ad-hoc and canned (pre-constructed) queries as well as generate reports.

The Tool You Should Use

Depending on your needs, you can access the Warehouse from your web browser or you can acquire and install the Cognos Suite on your computer. Your choice will depend on the type of user you are.

If you are what the Gartner Group calls a “typical” user, you don’t wish to invest a great deal of time learning technology or business intelligence applications. Generally 60 percent or more of people who will use the Warehouse fall into this category. You are considered a “typical” user if your needs for business intelligence are basic and straightforward and rarely change. If you’ve surfed the Internet, learning to use the Warehouse from your web browser will be fast and simple—little or no cost involved, except your time.

If you are what the industry calls a “dynamic” user, you wish to minimize your exposure to technology, but you also need up-to-the-minute data and the flexibility to reduce data to a more meaningful result set. For example, you want to filter out all the information you don’t need. Generally 20 to 30 percent of people who will use the Warehouse fall into this category. Once again, access to the Warehouse from your web browser is the method of choice, at little or no cost except time involved.

If you are a “free-form” user, you also need up-to-the-minute information. But, you need a great deal of flexibility in filtering and combining data, creating calculated (virtual) values and custom formatting for output. “Free-form” users are usually the ones who develop reports and queries that are later migrated to the web and used by “typical” and “dynamic” users. Acquiring and installing the Cognos Suite is the method of choice. The cost of the Cognos Suite is \$1,295 per person plus training costs. Training takes about a week and can be scheduled on site (\$3,500 per day, 3-day minimum, 12 student maximum) or at one of Cognos’ training sites (\$440 per day per student plus travel expenses).

Other Tools

Warehouse users may use any database query tool once they have received access to the Warehouse. Users of these tools are strongly advised to consult Data Administration before exploring these options for advice on connecting to the Warehouse, joining tables, and efficient querying.

Training

Department data is complex. Your success in executing queries and interpreting results depends on your having a working knowledge of department data, a basic understanding of how the Warehouse works, and a familiarity with using your web browser or the Cognos Suite to retrieve Warehouse data. Training will give you the framework you need to understand and use the Warehouse.

You must attend training for department data. If you will be a user of the desktop Cognos Suite, then training on this software is strongly recommended. You will be given access to the Warehouse only after you have attended the training that is appropriate for the data you are requesting. Contact any member of the Warehouse Team for information about department data and Cognos Suite training (www.hsdw.state.ut.us/support.htm).

Data training topics include the following. Additional information on training is also available for each subject area:

- Program terminology and concepts
- Subject areas, including detailed and summary data available

- How to use the data in your job
- How to create efficient queries
- Support services available after training
- How the Warehouse gets refreshed

Support Services

Technical Support

If you have a problem connecting to the Warehouse, using the query tool (for example, creating a query and formatting results), or understanding Warehouse data, contact any member of the Warehouse Team (www.hsdw.state.ut.us/support.htm).

Data Warehouse Mailing List

Upon receiving access to the Warehouse, we will automatically subscribe you to the Data Warehouse Mailing List. This electronic mail list announces the addition of new data in the Warehouse, changes in refresh cycles, changes in Warehouse availability, and other information of general interest. List members often discuss query methods and data questions via the lists as well.

If you are not receiving e-mail concerning general issues about the Warehouse or e-mail concerning a particular subject area, contact any member of the Warehouse Team to be added to the Data Warehouse Mailing List (www.hsdw.state.ut.us/support.htm).

Data Warehouse User Groups

Users of the Warehouse are periodically invited to attend group meetings to share experiences, ask questions, review data element use, and so on. In addition, smaller special interest groups may be formed to resolve specific issues. Warehouse users will be invited to meetings via the Data Warehouse Mailing List.

Warehouse Updates

Please refer to Notices (www.hsdw.state.ut.us/news/notices.htm) for specific information about Warehouse availability and refresh cycles.

In general, standard updates to the Warehouse take place on Sunday night. The refresh cycle varies for each subject area in the Warehouse. For example, client information is updated weekly, while provider information may be updated monthly. Other subject areas remain available while one subject area is being updated. Keep in mind, though, that overall query performance may be affected while updates are in process.

Reporting Data Quality Issues

We have made every attempt to assure that data in the Warehouse is accurate, such as minimizing duplicate client and service records that exist in source systems. When we become aware of data quality issues, we notify the people responsible for correcting problems with the data in each source system. However, it is impossible to achieve 100 percent accuracy. Keep quality of data in mind as you use the Warehouse. If you become aware of a

data quality problem, please contact any member of the Warehouse Team (www.hsdw.state.ut.us/support.htm).

What the Warehouse Contains

Client Information

Refresh Cycle

Client information is refreshed weekly on Sunday night, as are Services provided to clients. Related client tables are also updated during this time. (Related client tables include...)

Source of the Data

Client information that is accessible in the Warehouse comes from the mainframe USSDS files and mainframe JIS files.

Time Span Covered

Client information has been downloaded from mainframe USSDS files since its inception and continues from that point, up to the last refresh cycle. Client information has been downloaded from mainframe JIS files since January 1, 1975, for clients who were age five at that time, and continues from that point, up to the last refresh cycle. Refreshes generally include the preceding last two weeks of data. However, since both USSDS and JIS history can be changed at any time, we will periodically do a full load of data to capture changes in historical data.

Requesting Additions and Changes to the Warehouse

If you would like new data added to the Warehouse, or would like changes to be made to the Warehouse, please contact any member of the Warehouse Team (www.hsdw.state.ut.us/support.htm).